Dashboard for NFT

Ruohe Zhou
rzhou73@wisc.edu

Team Name: UWM

Tsz Yau Iris Chow
tchow7@wisc.edu

1. Introduction

Digital assets has become very popular over the years. NFT has been used to do auction in arts, music and sports. NFT is relatively new when compare to other cryptographic digital assets. There is a lot more to explore in terms of the stability and price potential in NFT. Therefore, our analysis aims to select variables to predict final price of NFT. Moreover, we visualised different variables to observe the relationship with final selling price.

2. Proposed Method

We used Shiny in R to build an interactive dashboard webpage and used Python to fit machine learning models to predict the final prices of digital products based on variables including auction type, last sale final price, favorites, duration.

2.1. Dataset

There are three tables given: event, asset and collection. For performing analysis, we merge tables to produce insights. We made a table that contains 'asset_token_id', 'collection_slug', 'created_date', 'id', 'quantity', 'listing_time', 'event_type', 'total_price', 'auction_type', 'duration', 'primary_asset_contracts_address', 'display_data_card_display_style', 'safelist_request_status', 'telegram_url', 'twitter_username', 'instagram_username', 'discord_url', 'medium_username', 'external_url', 'token_id', 'owner_address', 'last_sale_total_price', 'name', 'num_sales', 'asset_contract_owner', 'asset_favorites', 'asset_category', 'num_check', 'favorite', 'date'. Furthermore, we removed outliers and replaced NA values with the average of the column when performing analysis.

3. Results and Discussion

Shinyapp webpage: Click here to see webpage

Created event is the most frequent event type in the dataset. Successful event type and cancelled event type both are about 20 percent of the total.

99 percent of the event in the dataset are public, with only 1 percent of them to be private. (Figure 2)

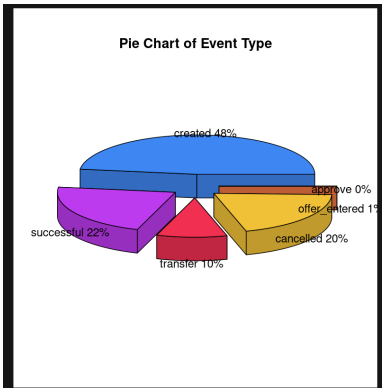

Figure 1. Event type Pie

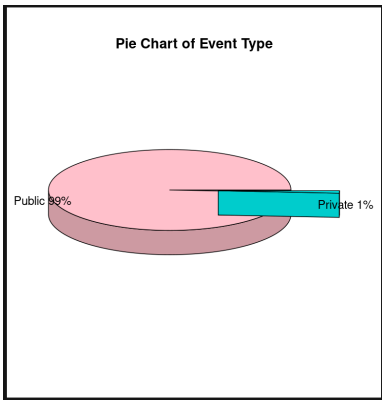

Figure 2. Private or Public Pie

99 percent of the event in the dataset are hosted in Dutch, with only 1 percent of them hosted by English. (Figure 3)

Accoring to Figure 4, we observed that collectibles, trading-cards, utility, and virtual-worlds these four types of asset category tend to receive more likes than others.

We also observed that the science and technology asset category has more concentrated range of event duration than other types of asset category. Art, Collectibles, and virtual-worlds each have one or two events that last a significant amount of time. (Figure 5)

According to the heat map(Figure 6), we observed that the correlations between each variables are very weak, which might explain why the prediction accuracy of linear

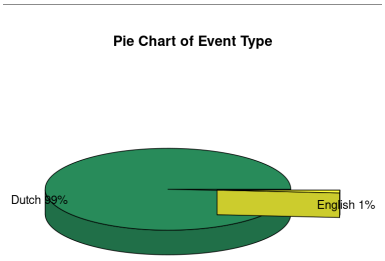

Figure 3. Auction Type Pie

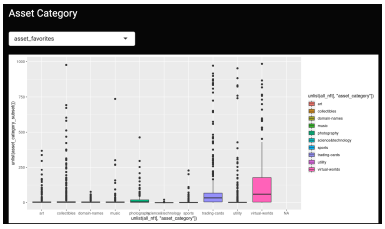

Figure 4. Event type Box

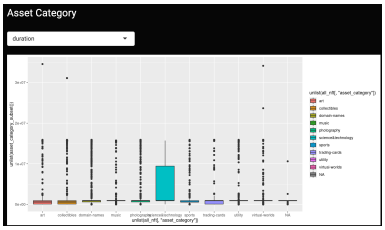

Figure 5. Duration Box

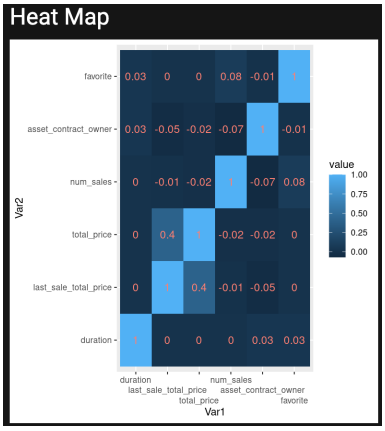

Figure 6. Heat map

regression is very low. The correlation between favorites (number of likes) and the number of sales is the highest (0.08).

As shown in bubble plot(Figure 7), favourites does not

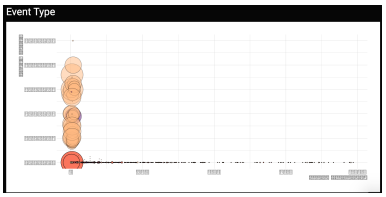

Figure 7. Bubble Plot

correlate with final price, and last sale final price and final price tend to follow the similar distribution pattern.

3.1. Machine Learning Models

We fit linear regression model using last sale total price as x variable and total price as y variable. We got an accuracy of 29%. Surprisingly, Linear Regression model and random forest model with multiple variables include the number of favorite, number of sales, duration, and last sale total price has lower accuracy with 15% and 14% respectively.

Table 1: Machine Learning Models

Method	Accuracy (R square)
Linear Regression with 1 variable	29%
Linear Regression with multiple variables	15%
Random Forest	14%

Table 1: Machine Learning Models